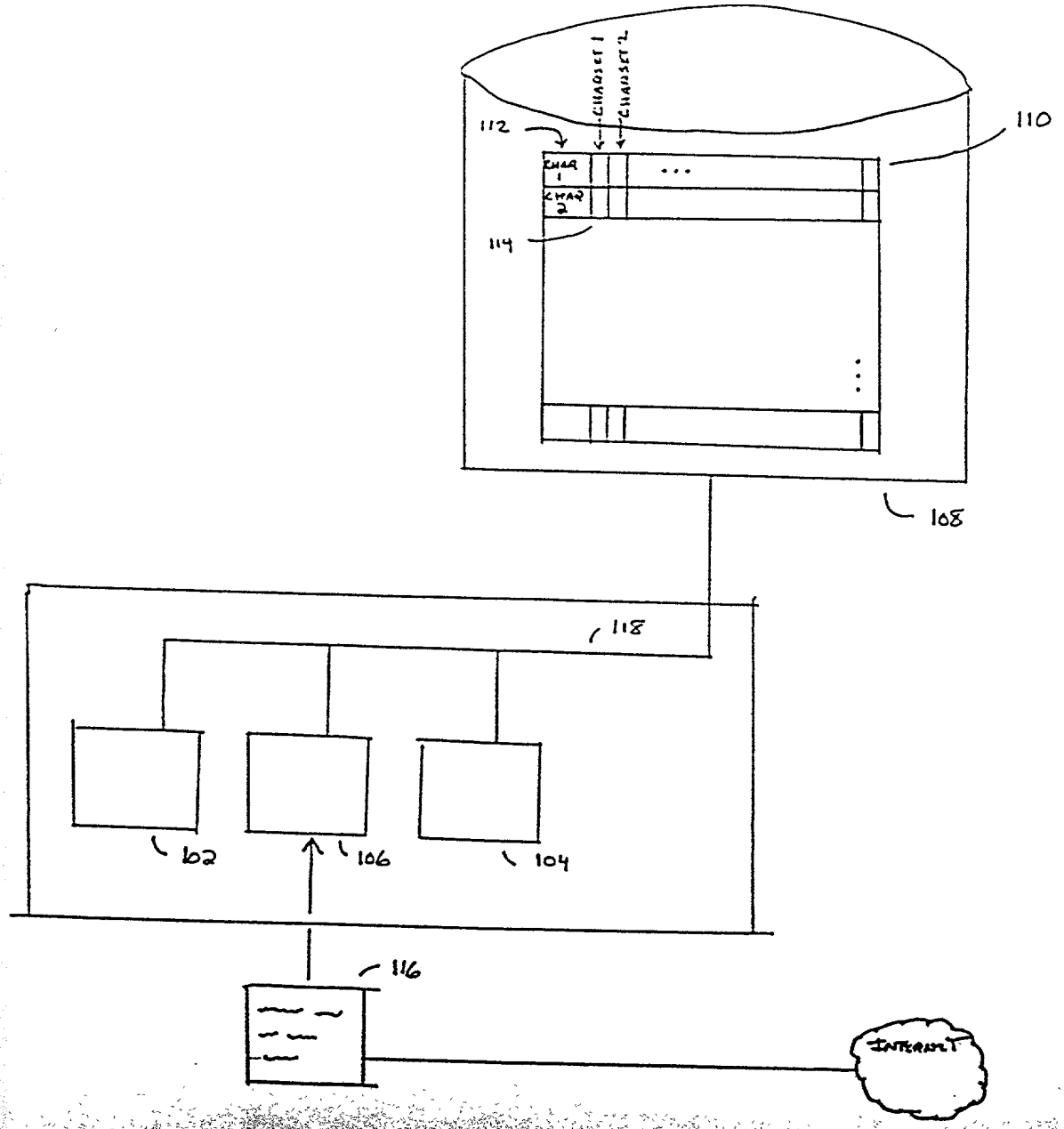

CHARSET 1
CHARSET 2
112
110
CHAR 1
CHAR 2
114
108
118
102
106
104
116
INTERNET

FIG. 1

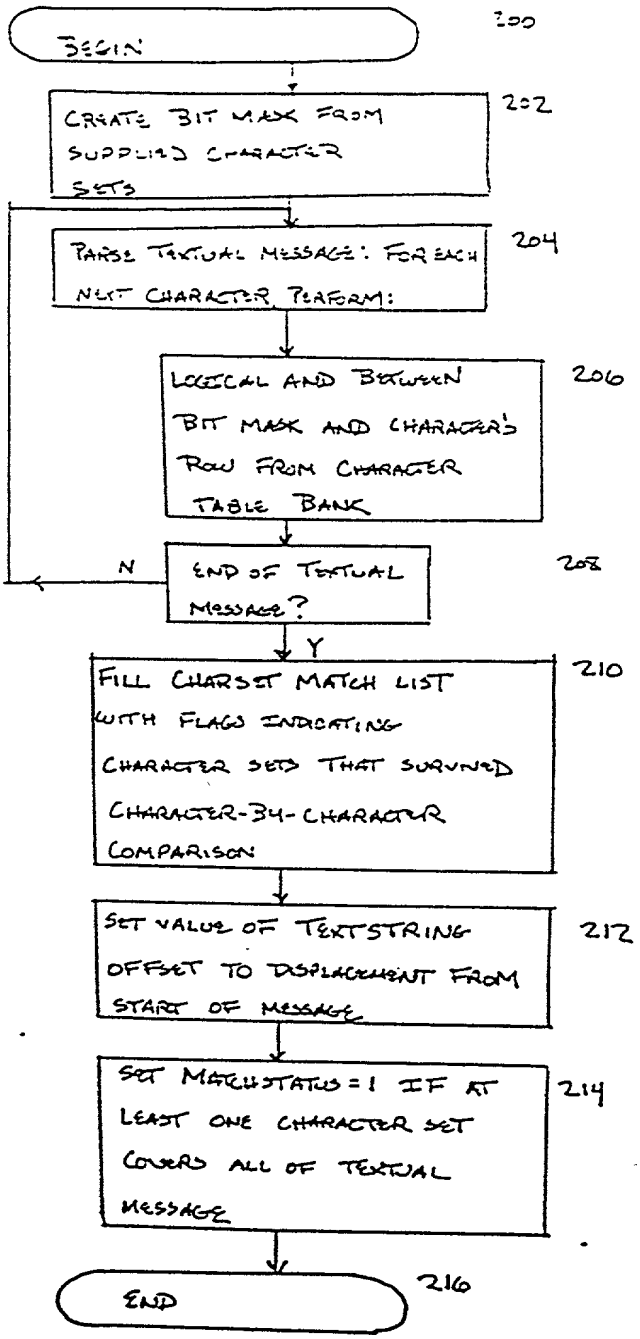

BEGIN
200
CREATE BIT MASK FROM SUPPLIED CHARACTER SETS
202
PARSE TEXTUAL MESSAGE: FOR EACH NEXT CHARACTER, PERFORM:
204
LOGICAL AND BETWEEN BIT MASK AND CHARACTER'S ROW FROM CHARACTER TABLE BANK
206
N
END OF TEXTUAL MESSAGE?
208
Y
FILL CHARSET MATCH LIST WITH FLAGS INDICATING CHARACTER SETS THAT SURVIVED CHARACTER-BY-CHARACTER COMPARISON
210
SET VALUE OF TEXTSTRING OFFSET TO DISPLACEMENT FROM START OF MESSAGE
212
SET MATCHSTATUS = 1 IF AT LEAST ONE CHARACTER SET COVERS ALL OF TEXTUAL MESSAGE
214
END
216

FIG. 2

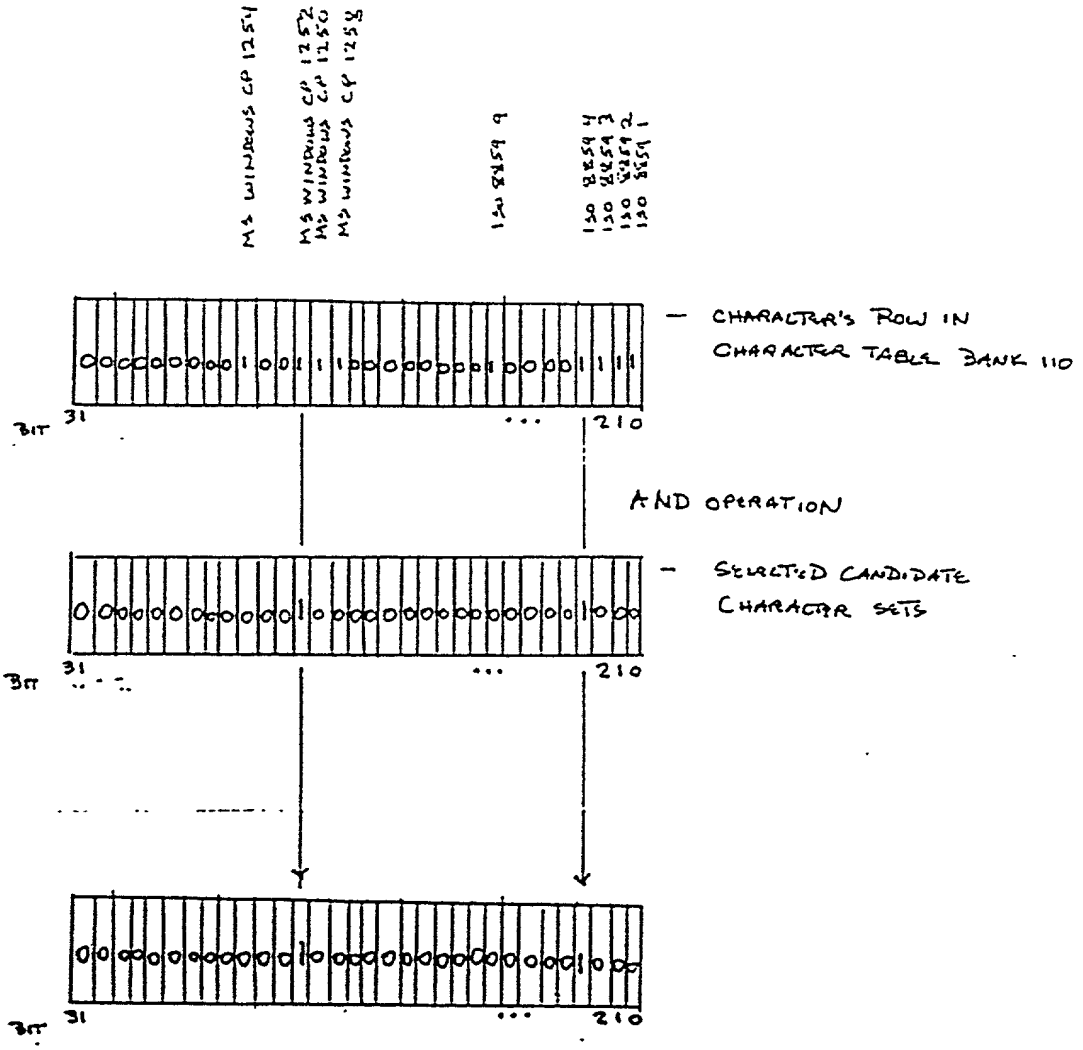
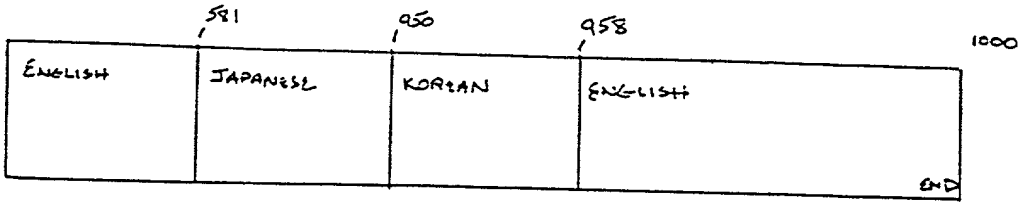

MS WINDOWS CP 1254
MS WINDOWS CP 1252
MS WINDOWS CP 1250
ISO 8859 9
ISO 8859 4
ISO 8859 3
ISO 8859 2
ISO 8859 1
CHARACTER'S ROW IN CHARACTER TABLE BANK 110
AND OPERATION
SELECTED CANDIDATE CHARACTER SETS
BIT
31
210

FIG. 3



1ST CALL: STOP AT 950

RET SET = JAPANESE
(CHARSETMATCHLIST)

MATCH STATUS = FAILED

2ND CALL: STOP AT 1000

RET SET = KOREAN
(CHARSETMATCHLIST)

MATCH STATUS = SUCCESS

FIG. 4

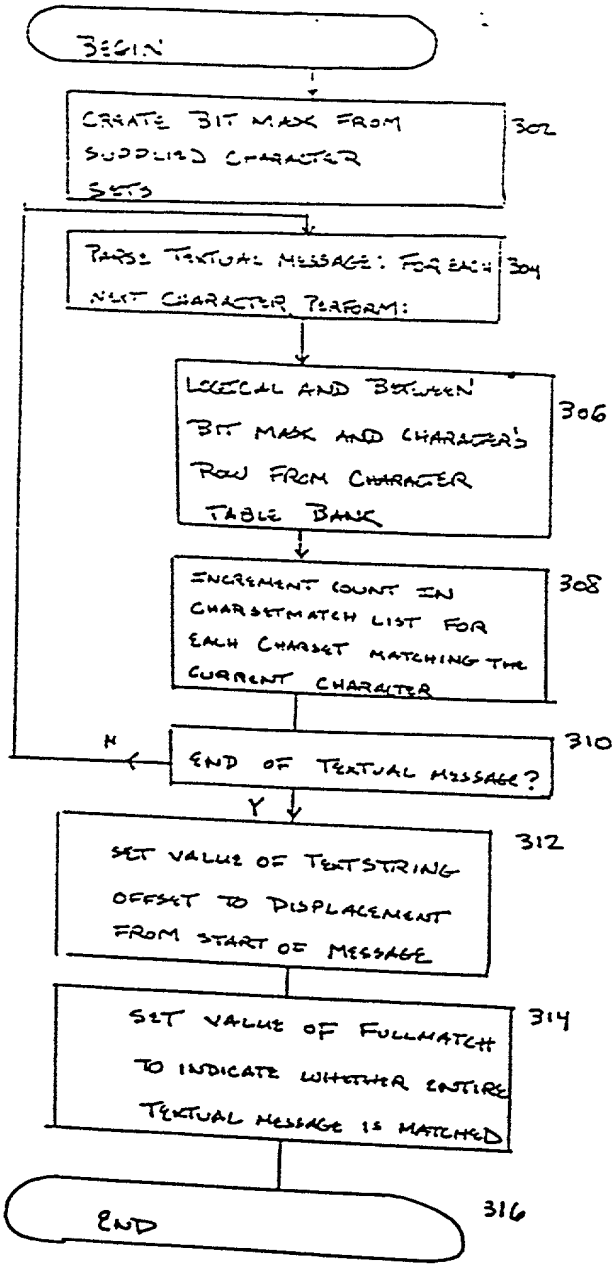

BEGIN
CREATE BIT MASK FROM SUPPLIED CHARACTER SETS
302
PARSE TEXTUAL MESSAGE: FOR EACH NEXT CHARACTER PERFORM:
304
LOGICAL AND BETWEEN BIT MASK AND CHARACTER'S ROW FROM CHARACTER TABLE BANK
306
INCREMENT COUNT IN CHARSETMATCH LIST FOR EACH CHARSET MATCHING THE CURRENT CHARACTER
308
N
END OF TEXTUAL MESSAGE?
310
Y
SET VALUE OF TEXTSTRING OFFSET TO DISPLACEMENT FROM START OF MESSAGE
312
SET VALUE OF FULLMATCH TO INDICATE WHETHER ENTIRE TEXTUAL MESSAGE IS MATCHED
314
END
316

FIG. 5

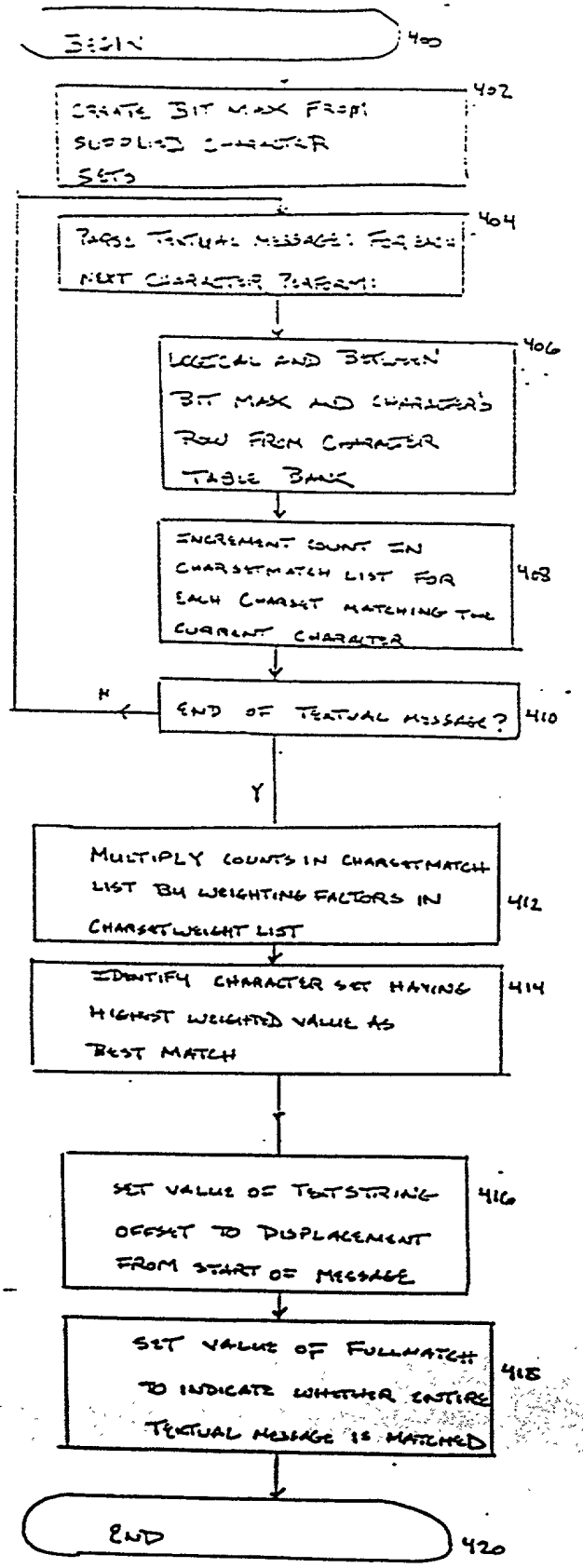
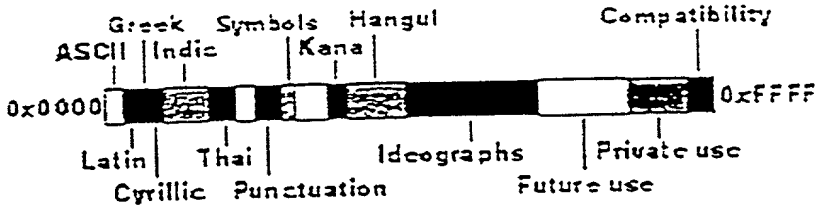

BEGIN
400
CREATE BIT MAX FROM SUPPLIED CHARACTER SETS
402
PARSE TEXTUAL MESSAGE: FOR EACH NEXT CHARACTER PERFORM:
404
LOCATE AND RETURN BIT MAX AND CHARACTER'S ROW FROM CHARACTER TABLE BANK
406
INCREMENT COUNT IN CHARSETMATCH LIST FOR EACH CHARSET MATCHING THE CURRENT CHARACTER
408
N
END OF TEXTUAL MESSAGE?
410
Y
MULTIPLY COUNTS IN CHARSETMATCH LIST BY WEIGHTING FACTORS IN CHARSETWEIGHT LIST
412
IDENTIFY CHARACTER SET HAVING HIGHEST WEIGHTED VALUE AS BEST MATCH
414
SET VALUE OF TESTSTRING OFFSET TO DISPLACEMENT FROM START OF MESSAGE
416
SET VALUE OF FULLMATCH TO INDICATE WHETHER ENTIRE TEXTUAL MESSAGE IS MATCHED
418
END
420

FIG. 6



Unicode's encoding layout

Fig. 7

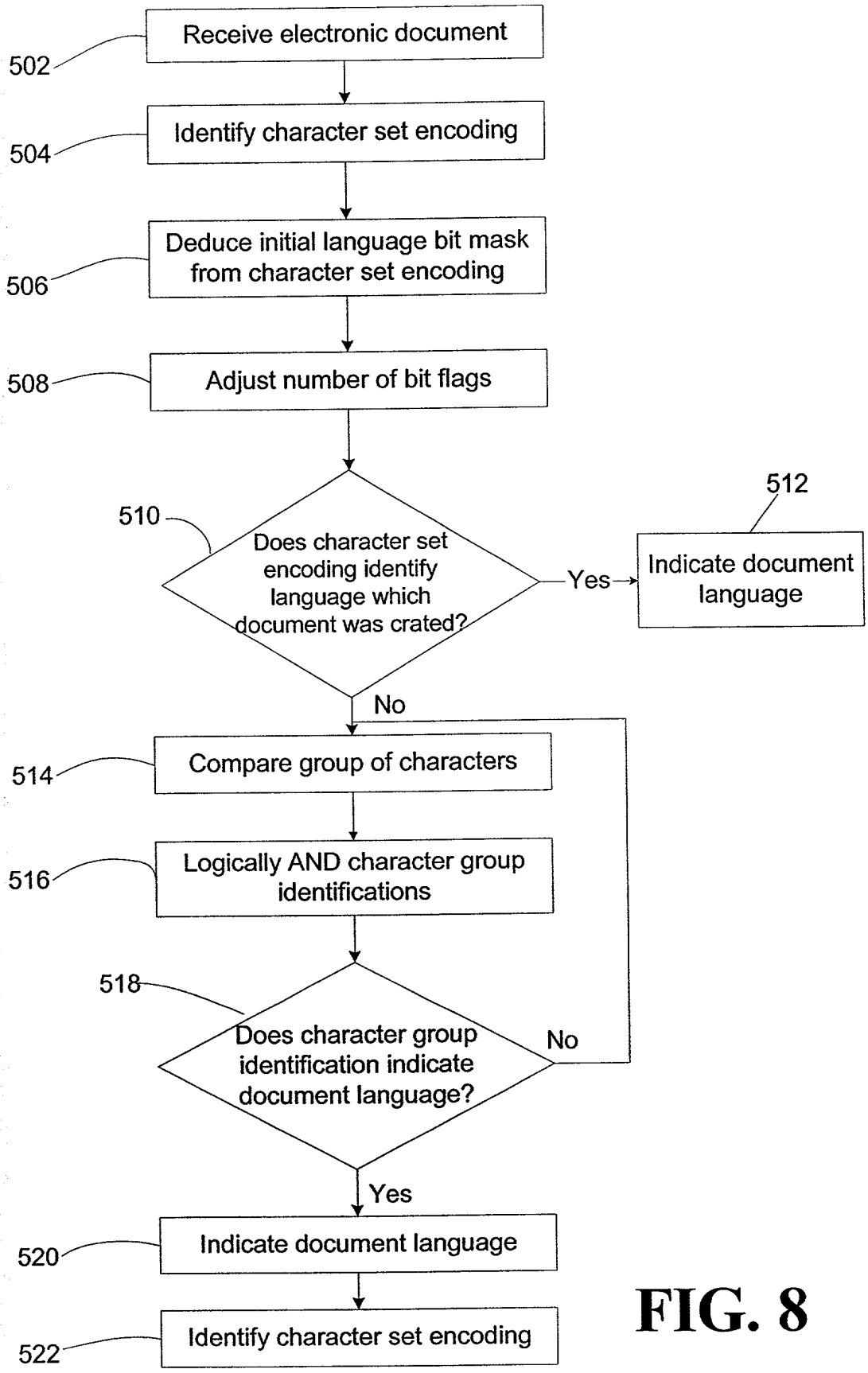

Receive electronic document
502
Identify character set encoding
504
Deduce initial language bit mask from character set encoding
506
Adjust number of bit flags
508
510
Does character set encoding identify language which document was crated?
Yes
512
Indicate document language
No
514
Compare group of characters
516
Logically AND character group identifications
518
Does character group identification indicate document language?
No
Yes
520
Indicate document language
522
Identify character set encoding

FIG. 8

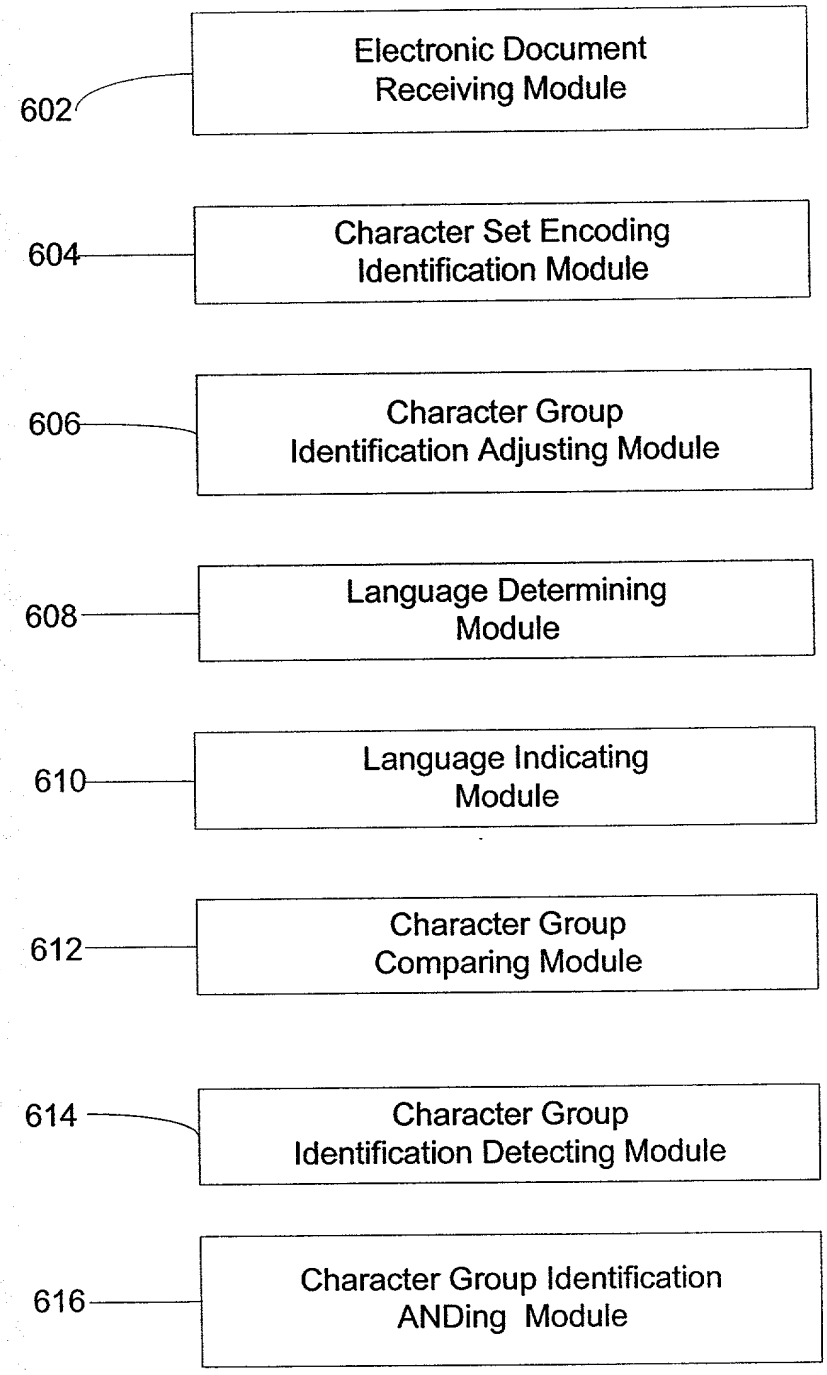

602
Electronic Document Receiving Module
604
Character Set Encoding Identification Module
606
Character Group Identification Adjusting Module
608
Language Determining Module
610
Language Indicating Module
612
Character Group Comparing Module
614
Character Group Identification Detecting Module
616
Character Group Identification ANDing Module

FIG. 9